

التقويم والقياس والاختبار والعلاقة بينهما

تأليف

مدرس الدكتور محمد لوتي

أكتوبر 1, 2025

اقتبس من هذا المقال

مدرس الدكتور محمد لوتي (2025). التقويم والقياس والاختبار والعلاقة بينهما. عرب سايكولوجي. تم الاسترجاع من <https://arabpsychology.com/?p=28073>

التقويم والقياس والاختبار: مفاهيم أساسية وعلاقات تكاملية في علم النفس والتربية

مقدمة: ركائز فهم الأداء البشري

في عالم يتزايد فيه الاهتمام بفهم قدرات الأفراد وتطوير مسارات التعلم والنمو البشري، تبرز ثلاثة مفاهيم محورية تشكل ركائز أساسية في مجالي علم النفس والتربية: الاختبار (Testing)، والقياس (Measurement)، والتقويم (Assessment). على الرغم من تمايزها الدقيق وأهميتها البالغة، غالباً ما تُستخدم هذه المصطلحات بشكل متبادل أو يُساء فهمها، مما قد يؤدي إلى ارتباك كبير في التطبيق العملي وفي اتخاذ القرارات المصيرية المتعلقة بالأفراد والبرامج التعليمية. إن التمييز الواضح بين هذه المفاهيم ليس مجرد تمرين أكاديمي، بل هو ضرورة عملية وحتمية لضمان فعالية التدخلات التعليمية ودقة الأحكام الصادرة حول قدرات الأفراد وسلامة تصميم البرامج التطويرية.

يهدف هذا المقال إلى تقديم تحليل شامل ومتعمق لكل من هذه المفاهيم، مبيناً تعريفاتها الجوهرية، وخصائصها المميزة، وأغراضها الأساسية التي تخدمها في السياقات التربوية والنفسية. وسيسعى جاهداً إلى كشف النقاب عن طبيعة العلاقة التكاملية والمتداخلة التي تجمع بين الاختبار والقياس والتقويم، موضحاً كيف يشكل كل منها لبنة أساسية لا يمكن الاستغناء عنها في صرح فهم وتقييم الأداء البشري وتطويره المستمر. من خلال فهم عميق لكل مصطلح على حدة، وكيفية تداخله وتكامله مع الآخرين، نتمكن من بناء أسس متينة لعمليات تقييم عادلة وموثوقة تسهم في تحقيق أفضل النتائج الممكنة للأفراد والمؤسسات على حد سواء.

الاختبار: الأداة المنظمة لجمع عينات السلوك

يُعد الاختبار، في أبسط تعريفاته، أداة أو إجراءً منظماً ومقنناً صُمم خصيصاً للحصول على عينة محددة من سلوك الفرد في مجال معين، ومن ثم تقييم هذه العينة في ضوء معايير محددة مسبقاً. يمكن النظر إليه على أنه "موقف مقنن يتضمن مجموعة من المثيرات، سواء كانت أسئلة أو مهام أو مشكلات، يُطلب من الفرد الاستجابة لها بطريقة معينة، بهدف الحصول على مؤشر كمي أو كيفي لسمه أو قدرة معينة" (Cohen & Swerdlik, 2018, p. 4). إنه بمثابة عدسة مكبرة تركز على جانب معين من جوانب الأداء البشري في لحظة زمنية محددة، مقدمة بذلك لقطة سريعة وموجهة.

تتميز الاختبارات الفعالة عادةً بمجموعة من الخصائص المنهجية التي تهدف إلى ضمان أكبر قدر ممكن من الموضوعية والدقة والموثوقية في عملية جمع البيانات. على سبيل المثال، تُعد الاختبارات المقننة نموذجاً بارزاً لهذه الخصائص؛ فهي تُطبق وتُصحح وتُفسر نتائجها بنفس الطريقة الموحدة لجميع الأفراد الذين يخضعون لها، مما يتيح مقارنة أداء الفرد بأداء مجموعة مرجعية كبيرة وممثلة (Norm Group). تتضمن عملية التقنين توحيد تعليمات الاختبار، وتحديد الزمن المخصص للإجابة بدقة، وتوحيد طريقة عرض الأسئلة، وتحديد إجراءات التصحيح التفصيلية، وكل ذلك يهدف إلى تقليل تأثير العوامل الخارجية غير المرتبطة بالقدرة المقاسة على أداء المفحوص.

تتنوع الاختبارات بشكل كبير لتخدم أغراضاً تعليمية ونفسية وتشخيصية متعددة، مما يعكس مرونتها وقدرتها على التكيف مع احتياجات التقييم المختلفة. فنجد اختبارات التحصيل التي تُستخدم لقياس مدى إتقان الفرد لمعارف أو مهارات محددة تم تعلمها مسبقاً، واختبارات الاستعداد التي تهدف إلى التنبؤ بقدرة الفرد على تعلم مهارة جديدة أو النجاح في مجال معين مستقبلاً. بالإضافة إلى ذلك، هناك اختبارات القدرات العقلية كاختبارات الذكاء، واختبارات الشخصية التي تسير أغوار سمات الفرد وأنماط سلوكه، واختبارات التشخيصية التي تركز على تحديد نقاط القوة والضعف المحددة لدى الفرد. يمكن أن تتخذ الاختبارات أشكالاً عديدة، فقد تكون فردية أو جماعية، ورقية أو محوسبة، محددة بوقت أو غير محددة، مما يوسع من نطاق استخدامها.

على الرغم من الأهمية الكبيرة للاختبارات كأدوات فعالة لجمع البيانات بسرعة وكفاءة نسبية، إلا أنه من الضروري إدراك أنها ليست سوى وسيلة واحدة ضمن منظومة تقييم أوسع وأشمل. فالاختبار يقدم لقطة لأداء الفرد في لحظة زمنية معينة وتحت ظروف محددة، وقد يتأثر هذا الأداء بمجموعة من العوامل غير المرتبطة بالسمة المقاسة، مثل قلق الاختبار أو الحالة الصحية. لذلك، فإن الاعتماد الكلي على نتيجة اختبار واحد لاتخاذ قرارات مهمة قد يكون مضللاً أو غير عادل. الاختبار في جوهره هو أداة لجمع البيانات، وليس غاية في حد ذاته، والمعلومة الأولية التي يقدمها تحتاج إلى سياق أعمق وتفسير دقيق لتصبح ذات معنى حقيقي وقيمة، وهذا ما يقودنا إلى المفهوم الثاني: القياس.

القياس: تحويل السلوك إلى بيانات كمية ذات معنى

إذا كان الاختبار هو الأداة المنهجية التي نستخدمها لجمع البيانات، فإن القياس هو العملية الجوهرية التي يتم من خلالها توظيف هذه الأداة (أو أدوات أخرى متنوعة) لتعيين قيمة عددية أو رمزية لخاصية أو سمة معينة لدى الفرد، وذلك وفقاً لقواعد محددة وموضوعية. يُعرف إدوارد ثورندايك (Thorndike)، أحد الرواد الأوائل في مجال القياس النفسي والتربوي، القياس بأنه "عملية تحديد مقدار ما يوجد في الشيء من الخاصية المراد قياسها" (Thorndike, 1904). بشكل أساسي، يسعى القياس إلى الإجابة عن أسئلة كمية مثل "كم؟" أو "إلى أي مدى؟" أو "ما هو مقدار؟". إنه يمثل عملية تحويل الملاحظات النوعية أو الخصائص المجردة، مثل المعرفة بالرياضيات أو مستويات القلق، إلى بيانات كمية قابلة للمعالجة الإحصائية والتحليل والمقارنة عبر الأفراد أو عبر الزمن.

إن القواعد التي يتم بموجبها تعيين الأرقام هي جوهر عملية القياس وروحها النابضة. هذه القواعد يجب أن تكون واضحة، ومحددة، ومتسقة لضمان أن الرقم المُعين يعكس بالفعل، وبأكبر قدر ممكن من الدقة، درجة وجود السمة التي نسعى لقياسها لدى الفرد. على سبيل المثال، في اختبار تحصيلي، قد تُحدد القاعدة بأن "تُعطي درجة واحدة لكل إجابة صحيحة، وصفر للإجابات الخاطئة". وفي مقياس ليكرت، قد تكون القاعدة هي "تعيين الرقم 5 للموافقة بشدة، 4 للموافقة، وهكذا". تعتمد دقة القياس وفائدته بشكل حاسم على جودة الأداة المستخدمة ومدى ملاءمة ووضوح القواعد المتبعة لتعيين الأرقام وتفسيرها، فضلاً عن مدى صدق هذه الأداة.

من الأهمية بمكان إدراك أن طبيعة القياس في العلوم التربوية والنفسية تختلف جوهرياً عن القياس في العلوم الفيزيائية والطبيعية، وهو فرق يستدعي تفكيراً متأنياً. فبينما يمكننا قياس الطول أو الوزن بوحدات ثابتة عالمياً وذات صفر مطلق، فإن السمات النفسية والتربوية، مثل الذكاء أو التحصيل الأكاديمي، هي في الغالب مفاهيم مجردة (constructs) لا يمكن ملاحظتها مباشرة. نحن نقيسها بشكل غير مباشر من خلال ملاحظة السلوكيات الظاهرة أو الاستجابات التي يُفترض أنها تعكس هذه السمات. لتوضيح هذا التباين، قام ستانلي سميث ستيفنز (Stevens, 1946) بتصنيف مستويات القياس إلى أربعة أنواع رئيسية، والتي توضح بدقة طبيعة الأرقام والمعلومات التي يمكن استخلاصها منها:

المقياس الاسمي (Nominal Scale): تُستخدم الأرقام كمجرد تسميات أو رموز لتمييز وتصنيف الأشياء أو الأفراد في فئات مختلفة دون أي دلالة على مقدار أو ترتيب. هذه الأرقام لا يمكن ترتيبها أو إجراء عمليات حسابية عليها.

المقياس الترتيبي (Ordinal Scale): تعكس الأرقام في هذا المقياس ترتيباً أو رتبة للأفراد أو الأشياء وفقاً لدرجة امتلاكهم للسمة المقاسة. لكن المسافات أو الفروق بين الرتب المتتالية ليست بالضرورة متساوية أو ذات معنى كمي متساوٍ.

المقياس الفئوي أو الفاصل (Interval Scale): الأرقام في هذا المستوى لا تعكس الترتيب فحسب، بل تكون المسافات بينها متساوية وذات معنى كمي متسق. ومع ذلك، فإن الصفر في هذا المقياس هو صفر افتراضي أو اصطلاحية، ولا يعني انعدام السمة تماماً.

المقياس النسبي (Ratio Scale): يتمتع هذا المقياس بجميع خصائص المقاييس السابقة، بالإضافة إلى وجود صفر مطلق حقيقي يعني انعدام السمة تماماً. يسمح هذا المقياس بإجراء جميع العمليات الحسابية وتفسير النسب بين القيم.

إن معظم القياسات في مجالي التربية وعلم النفس تقع ضمن المستويين الترتيبي والفتوي (Bond & Fox, 2015). إن فهم مستوى القياس الذي تنتمي إليه البيانات التي نجمعها أمر بالغ الأهمية؛ لأنه يحدد بوضوح نوع التحليلات الإحصائية التي يمكن إجراؤها على هذه البيانات ومدى مشروعية التفسيرات التي يمكن استخلاصها منها. القياس، إذن، يوفر لنا البيانات الكمية الضرورية، ولكنه لا يزال لا يقدم الحكم القيمي أو القرار النهائي. إنه يصف لنا الوضع الحالي بناءً على الأداة المستخدمة، لكن الحكم على جودة هذا الوضع أو اتخاذ قرار بشأنه يتطلب الانتقال إلى المفهوم الأوسع والأكثر شمولاً، وهو التقويم.

التقويم: العملية الشاملة للحكم واتخاذ القرار

يُعد التقويم المفهوم الأكثر شمولاً وعمومية بين المفاهيم الثلاثة، ويمثل القمة في هذه المنظومة المتكاملة. إذا كان الاختبار هو الأداة المحددة لجمع البيانات، والقياس هو العملية المنهجية لإعطاء قيمة رقمية لهذه البيانات، فإن التقويم هو "عملية منظمة ومتعددة الأوجه لجمع المعلومات حول تعلم الطلاب أو أدائهم أو تطورهم، ثم تحليل هذه المعلومات وتفسيرها بدقة، بهدف إصدار أحكام قيمية مستنيرة أو اتخاذ قرارات هادفة" (Linn & Gronlund, 2000, p. 31). التقويم لا يقتصر بالضرورة على استخدام الاختبارات أو الحصول على بيانات كمية فقط، بل يتجاوز ذلك ليشمل جمع طيف واسع ومتنوع من الأدلة، التي يمكن أن تكون كمية وكيفية، ومن مصادر متعددة ومتنوعة.

تشمل مصادر المعلومات المستخدمة في التقويم مجموعة واسعة من الأدوات والتقنيات، منها الاختبارات بأنواعها المختلفة، والملاحظات الصفية المنظمة، ومراجعة أعمال الطلاب (مثل المشاريع والواجبات اليومية)، وملفات الإنجاز (البورتفوليو) التي توثق تطور الطالب على مدار فترة زمنية، والمقابلات الشخصية مع الطلاب وأولياء الأمور، وتقييمات الأقران، والتقييم الذاتي من قبل الطالب نفسه. كل هذه المصادر تُجمع لتقديم صورة متكاملة وشاملة عن الفرد أو عن البرنامج التعليمي أو التدريبي قيد الدراسة (Popham, 2017). إن هذه النظرة الشمولية هي ما يميز التقويم عن مجرد الاختبار أو القياس.

السمة الجوهرية التي تميز التقويم بشكل قاطع عن القياس هي عنصر "الحكم القيمي" وعملية "اتخاذ القرار". فالتقويم لا يكتفي بوصف الأداء كما يفعل القياس؛ على سبيل المثال، لا يكتفي بالقول "حصل الطالب على 70% في الاختبار" (وهو قياس)، بل يتجاوزوه إلى تفسير معنى هذا الأداء في ضوء أهداف تعليمية محددة مسبقاً أو معايير أداء معينة. فالتقويم قد يقول: "أداء الطالب جيد ويظهر فهماً لمعظم المفاهيم الأساسية، لكنه بحاجة إلى دعم إضافي في حل المشكلات المعقدة التي تتطلب تفكيراً نقدياً". يهدف التقويم في جوهره إلى التحسين المستمر، سواء كان ذلك تحسين تعلم الطلاب، أو تطوير المناهج الدراسية لتصبح أكثر فعالية، أو تقييم مدى نجاح برنامج تدريبي في تحقيق أهدافه، أو اتخاذ قرارات إدارية مهمة مثل تحديد مستوى الطلاب، أو منح الشهادات، أو التوجيه المهني.

يمكن تصنيف التقويم بناءً على الغرض الأساسي الذي يسعى لتحقيقه إلى أنواع رئيسية متعددة، كل منها يخدم هدفاً محدداً في العملية التعليمية والنفسية:

التقويم التكويني أو البنائي (Formative Assessment): يحدث بشكل مستمر أثناء عملية التعلم والتعليم. هدفه الأساسي هو مراقبة تقدم الطلاب وتقديم تغذية راجعة فورية ومستمرة لكل من الطالب والمعلم لدعم عملية التعلم وتوجيهها.

التقويم الختامي أو التجميعي (Summative Assessment): يحدث عادةً في نهاية فترة تعليمية معينة. هدفه هو تقييم مدى

تحقيق الطلاب للأهداف التعليمية المحددة مسبقاً وإصدار حكم نهائي على مستوى أدائهم أو كفاءتهم الشاملة، وغالباً ما يستخدم لأغراض منح الدرجات أو إصدار الشهادات.

التقويم التشخيصي (Diagnostic Assessment): يهدف إلى تحديد نقاط القوة والضعف المحددة لدى الطالب قبل بدء عملية التدريس أو في بدايتها، أو لتحديد الأسباب الكامنة وراء صعوبات التعلم، مما يساعد المعلم على تخطيط تدخلات علاجية مناسبة.

التقويم الأصيل أو الواقعي (Authentic Assessment): يركز على تقييم قدرة الطلاب على تطبيق معارفهم ومهاراتهم في مواقف حياتية حقيقية أو محاكية للواقع بشكل وثيق، ويتضمن مهام معقدة تتطلب التفكير النقدي وحل المشكلات.

يتضح من هذا التنوع الشديد أن التقويم عملية ديناميكية، معقدة، ومتعددة الأوجه. إنه ليس مجرد حدث منفرد أو إجراء روتيني، بل هو عملية مستمرة ومتكاملة في الممارسة التربوية والنفسية الفعالة. يعتمد التقويم الجيد والفعال على استخدام أدوات قياس موثوقة وصادقة، ولكنه لا يقتصر عليها بأي حال من الأحوال. بل يشمل تفسير البيانات التي تم جمعها في سياقها الأوسع، واستخدامها لإصدار أحكام مستنيرة وذات قيمة تخدم غرضاً تعليمياً أو تطويرياً محدداً، مما يجعله المحرك الأساسي للتحسين والنمو.

العلاقة التكاملية بين الاختبار والقياس والتقويم

بعد أن استعرضنا المفاهيم الثلاثة – الاختبار، والقياس، والتقويم – بشكل منفصل ومتعمق، تتضح الآن بشكل جلي طبيعة العلاقة العضوية والتكاملية التي تربط بينها. لا يمكن فصل هذه المفاهيم عن بعضها البعض؛ بل هي تشكل منظومة متناغمة ومتداخلة، يمكن تصورها على شكل هرمي أو علاقة احتواء منطقية: الاختبارات هي أدوات محددة تُستخدم غالباً في عملية القياس، والقياس بدوره هو عملية توفير البيانات (الكمية غالباً) التي تُستخدم كأحد المدخلات الأساسية في عملية التقويم الأوسع والأشمل. ببساطة، الاختبار جزء من القياس، والقياس جزء من التقويم. هذه العلاقة تنتمي إلى مجال القياس النفسي وعلم النفس التربوي، وهما فرعان أساسيان يسعيان إلى فهم وتطوير أدوات وطرق تقييم الأداء البشري.

لتوضيح هذه العلاقة المعقدة والمتشابكة بشكل ملموس، دعونا نضرب مثلاً من خارج السياق التربوي، مستلهمين من مجال الطب. تخيل طبيباً يسعى لتقييم الحالة الصحية لمريض يشكو من أعراض معينة:

قد يستخدم الطبيب اختباراً محدداً، وهو مقياس الحرارة (Thermometer)، للحصول على قراءة دقيقة. مقياس الحرارة هنا هو الأداة.

القراءة التي يحصل عليها الطبيب من مقياس الحرارة، مثلاً "39 درجة مئوية"، تمثل قياساً دقيقاً لدرجة حرارة جسم المريض. هذا القياس هو بيان كمي موضوعي لخاصية فيزيائية، يجيب على سؤال "كم هي درجة الحرارة؟".

لكن الطبيب لا يتوقف عند هذا الرقم المجرد. إنه يقوم بعملية تقويم شاملة. يضع القياس (39 درجة) في سياقه الطبي: هل هذه الدرجة طبيعية لهذا العمر أو الحالة؟ ما هي الأعراض الأخرى المصاحبة؟ ما هو تاريخ المريض الطبي؟ بناءً على كل هذه المعلومات المتنوعة (الكمية من القياسات والكيفية من الأعراض والتاريخ)، يصدر الطبيب حكماً قيمياً ("هذا المريض يعاني من حمى مرتفعة وقد يكون لديه التهاب معين")، ثم يتخذ قراراً مستنيراً ("يحتاج المريض إلى خفض الحرارة وربما مضاد حيوي بعد فحوصات إضافية").

هذا المثال يوضح بوضوح أن القياس (قراءة درجة الحرارة) هو جزء ضروري ولا غنى عنه من عملية التقويم الشاملة، وأن الاختبار (مقياس الحرارة) هو الأداة المحددة للحصول على هذا القياس. لكن التقويم هو العملية الكلية التي تتجاوز مجرد

جمع الأرقام لتشمل جمع معلومات متنوعة، وتفسيرها في ضوء معايير وخبرات، وإصدار حكم ذي معنى، واتخاذ إجراءات بناءً على هذا الحكم.

في السياق التربوي، تتجلى العلاقة بين هذه المفاهيم بشكل مماثل:

قد يعطي المعلم اختباراً تحصيلياً (مجموعة أسئلة منظمة) في مادة الرياضيات لتقييم فهم الطلاب لوحدة دراسية معينة.

تصحيح الاختبار وإعطاء درجة كمية لكل طالب (مثلاً، 8 من 10) يمثل عملية قياس دقيقة لمستوى تحصيل الطالب في الموضوع الذي يغطيه الاختبار. هذا القياس يقدم لنا رقماً يدل على مستوى الأداء.

يقوم المعلم بعد ذلك بعملية تقويم متكاملة. فهو ينظر إلى درجة القياس (8/10)، ويقارنها بأداء الطالب السابق، وبأداء بقية الطلاب في الفصل، وبأهداف التعلم المحددة مسبقاً. قد يراجع إجابات الطالب ليرى أين أخطأ تحديداً ولماذا، هل كانت أخطاء مفاهيمية أم حسابية؟ قد يلاحظ سلوك الطالب أثناء الحصة، ومدى مشاركته، أو مستوى اهتمامه. قد ينظر أيضاً في واجباته المنزلية الأخرى. بناءً على كل هذه المعلومات المتنوعة، يصدر المعلم حكماً قيمياً ("فهم الطالب جيد للمفهوم العام ولكنه يواجه صعوبة في العمليات الحسابية المعقدة")، ثم يتخذ قراراً تربوياً ("سأعطيه تمارين إضافية مركزة على العمليات الحسابية، وسأراقبه عن كثب في الدرس القادم، وقد أنظم جلسة دعم فردية له").

أهمية التمييز بين المفاهيم وتأثيرها

إن العلاقة بين هذه المفاهيم هي علاقة تكاملية وليست تنافسية أو تبادلية. التقويم الفعال يعتمد بشكل أساسي على بيانات قياس دقيقة وموثوقة، والتي بدورها غالباً ما تتطلب استخدام اختبارات (أو أدوات قياس أخرى) جيدة التصميم، موثوقة، وصادقة ومناسبة للغرض المحدد. لا يمكن إجراء تقويم هادف وموضوعي بدون معلومات وبيانات موثوقة (وهذا هو دور القياس). كما لا يمكن الحصول على بيانات دقيقة وموضوعية في كثير من الأحيان بدون أدوات مصممة بعناية ومقننة (وهذا هو دور الاختبارات). بالمقابل، فإن مجرد إجراء الاختبارات والحصول على القياسات لا يكفي بحد ذاته؛ فالبيانات تصبح ذات معنى وقيمة حقيقية فقط عندما تُستخدم ضمن إطار تقويمي شامل يهدف إلى فهم أعمق للظاهرة المدروسة وتحقيق تحسين مستمر (Stiggins, 2005).

إن الخلط بين هذه المفاهيم أو استخدامها بشكل خاطئ قد يؤدي إلى ممارسات تقييمية غير فعالة أو حتى ضارة. فمثلاً، قد يركز المعلمون بشكل مفرط على الاختبارات كالمقياس الوحيد للنجاح (وهذا يمثل خلطاً بين الاختبار والتقويم)، مما قد يشجع على التعلم السطحي والحفظ البصم بدلاً من الفهم العميق وتنمية المهارات. أو قد يتم اتخاذ قرارات حاسمة ومصيرية بناءً على قياس واحد غير مكتمل أو غير كافٍ (وهذا يمثل خلطاً بين القياس والتقويم)، مما يؤدي إلى أحكام غير عادلة أو غير دقيقة.

إن الفهم الواضح والراسخ للعلاقة بين هذه المفاهيم الثلاثة يساعد المربين والممارسين على استخدام كل منها في مكانه الصحيح، وبالطريقة الأنسب، وبالتوقيت المناسب لتحقيق الأهداف المرجوة من عملية التقييم، مما يضمن تقيماً شاملاً وعادلاً يسهم في النمو والارتقاء. هذا الفهم يمثل حجر الزاوية في ممارسات التقييم المهنية، ويسهم في بناء ثقافة تعليمية ونفسية أكثر فعالية ومسؤولية.

الخاتمة: نحو ممارسات تقييمية مستنيرة

في ختام هذا الاستعراض المتعمق، يتضح أن كلاً من الاختبار، والقياس، والتقويم يمثل مكونات أساسية ومتراصة في صلب الممارسات التربوية والنفسية الحديثة والفعالة. فالاختبار يوفر لنا الأداة المنهجية والمنظمة لجمع عينات محددة من

السلوك أو الأداء، مما يتيح لنا الحصول على بيانات أولية. بينما يتولى القياس مهمة ترجمة هذه العينات السلوكية إلى بيانات كمية أو رمزية منظمة، قابلة للتحليل والتفسير، مجيباً على سؤال "كم؟" أو "إلى أي مدى؟". أما التقويم، فهو المفهوم الأشمل والأكثر تكاملاً، حيث يشمل العملية الكاملة لجمع وتفسير وتحليل مجموعة متنوعة من الأدلة (بما في ذلك بيانات القياس المستخلصة من الاختبارات) بهدف إصدار أحكام قيمية مدروسة واتخاذ قرارات مستنيرة وهادفة.

إن العلاقة بين هذه المفاهيم هي علاقة تداخل وتكامل هرمي لا يمكن فك عراها؛ حيث يعتبر الاختبار أداة محورية ضمن عملية القياس، بينما يُعد القياس خطوة أساسية لا غنى عنها ضمن عملية التقويم الأوسع نطاقاً والأكثر شمولية. كل مفهوم يخدم غرضاً فريداً، ولكن قيمته الحقيقية تتجلى عندما يتم توظيفه ضمن هذا الإطار المتكامل، حيث يدعم كل جزء الآخر ويقويه، مما يؤدي إلى نتائج أكثر دقة وموثوقية، ويسهم في بناء فهم شامل ومتكامل للظواهر النفسية والتربوية.

في نهاية المطاف، إن الفهم الدقيق لهذه المفاهيم العميقة والعلاقات الجوهرية التي تربطها ليس مجرد رفاهية أكاديمية، بل هو شرط أساسي وحتمي لضمان جودة وفعالية الممارسات التقييمية في مختلف المجالات، سواء كانت تربوية، نفسية، مهنية، أو حتى صحية. فمن خلال الاختيار الواعي لأدوات الاختبار المناسبة والأكثر ملاءمة للغرض، وتطبيق إجراءات القياس بدقة وموضوعية ووفقاً لمعايير علمية صارمة، وتوظيف نتائج القياس ضمن إطار تقويمي شامل ومتعدد المصادر، يمكن للمربين والأخصائيين النفسيين والمدربين الحصول على فهم أعمق وأكثر صدقاً لقدرات الأفراد واحتياجاتهم الحقيقية، مما يمهّد الطريق لاتخاذ قرارات أكثر حكمة وفعالية في جميع جوانب النمو البشري.

المراجع

- American Psychological Association. (2020). **Publication manual of the American Psychological Association** (7th ed.). <https://doi.org/10.1037/0000165-000>
- .Anastasi, A., & Urbina, S. (1997). **Psychological testing** (7th ed.). Prentice Hall
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. **Assessment in Education: Principles, Policy & Practice**, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Bond, T. G., & Fox, C. M. (2015). **Applying the Rasch model: Fundamental measurement in the human sciences** (3rd ed.). Routledge
- Cohen, R. J., & Swerdlik, M. E. (2018). **Psychological testing and assessment: An introduction to tests and measurement** (9th ed.). McGraw-Hill Education
- .Gronlund, N. E., & Waugh, C. K. (2009). **Assessment of student achievement** (9th ed.). Pearson
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), **Assessment and learning** (pp. 103-118). Sage Publications Ltd. <https://doi.org/10.4135/9781446211001.n7>
- Kaplan, R. M., & Saccuzzo, D. P. (2017). **Psychological testing: Principles, applications, and issues** (9th ed.). Cengage Learning
- .Linn, R. L., & Gronlund, N. E. (2000). **Measurement and assessment in teaching** (8th ed.). Prentice-Hall

.Nitko, A. J., & Brookhart, S. M. (2011). **Educational assessment of students** (6th ed.). Pearson

.Popham, W. J. (2017). **Classroom assessment: What teachers need to know** (8th ed.). Pearson

Stevens, S. S. (1946). On the theory of scales of measurement. **Science**, 103(2684), 677-680.
<https://doi.org/10.1126/science.103.2684.677>

Stiggins, R. J. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. **Phi Delta Kappan**, 87(4), 324-328.
<https://doi.org/10.1177/003172170508700414>

Wiggins, G. (1990). The case for authentic assessment. **Practical Assessment, Research & Evaluation**, 2(2).
<https://doi.org/10.7275/ffb1-mm19>

ARABPSYCHOLOGY.COM